

(11)特許出願公開番号

(43)公開日 平成10年(1998)10月27日

C

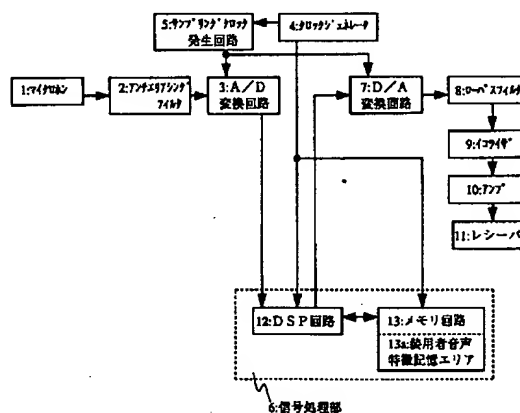
(71)出願人 000004075
ヤマハ株式会社
静岡県浜松市中沢町10番1号

(72)発明者 近藤 克文
静岡県浜松市中沢町10番1号 ヤマハ株式
会社内

(72)発明者 谷高 幸司
静岡県浜松市中沢町10番1号 ヤマハ株式
会社内

(74)代理人 弁理士 小森 久夫

【解決手段】マイクロホン１で装用者の音声信号を取り込み、その特徴を抽出して装用者音声特徴記憶エリア１３ａに記憶する。その後、マイクロホン１から音声信号が入力されると、その音声信号が装用者のものであるか否かを装用者音声特徴記憶エリア１３ａに記憶されている特徴に基づいて判断し、装用者が発声した音声信号の場合には、この音声信号を話速変換処理を行わないで出力する。一方、装用者以外が発声した音声信号の場合には、この音声信号に対して話速変換処理を行ったのち出力する。



【特許請求の範囲】

【請求項1】 取り込んだ音声信号を時間軸上で伸長する話速変換手段を備えた補聴器において、装用者の音声信号の特徴情報を記憶する特徴記憶手段と、該特徴情報に基づいて取り込んだ音声信号が装用者の発した音声であるかどうかを識別する音声識別手段と、前記音声識別手段で装用者の音声であることを識別したときには、前記話速変換手段を禁止する話速変換禁止手段と、を備えたことを特徴とする話速変換機能を有する補聴器。

【請求項2】 取り込んだ音声信号を時間軸上で伸長する話速変換手段を備えた補聴器において、装用者の音声信号を入力し、該音声信号の特徴情報を抽出する特徴抽出手段と、抽出された特徴情報を記憶する特徴記憶手段と、該特徴情報に基づいて取り込んだ音声信号が装用者の発した音声であるかどうかを識別する音声識別手段と、前記音声識別手段で装用者の音声であることを識別したときには、前記話速変換手段を禁止する話速変換禁止手段と、を備えたことを特徴とする話速変換機能を有する補聴器。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】この発明は、高齢化等により衰えた聴覚機能を補助する補聴器に関し、特に取り込んだ音声信号を時間軸上で伸長する話速変換処理を行って、装用者にゆっくり、はっきり聴かせる話速変換機能を有する補聴器に関する。

【0002】

【従来の技術】従来より、聴力が低下した人物の聴覚機能を補助するものとして補聴器が使用されている。ところで、聴力の低下が高齢化による場合等では、最小可聴信号レベルの上昇や高音域の聴取機能の低下のみならず音声識別臨界速度（音声を識別することができる最大の話速）も低下しており、このような人物には、単に音声の増幅や周波数特性のみ加工するだけではなく、時間特性も同時に加工することが望ましい。このため、最近では取り込んだ音声信号を時間軸上で伸長する話速変換処理を行って、この補聴器の装用者にゆっくり、はっきり聴かせることができる補聴器が考えられている。

【0003】音声信号を時間軸上で伸長させる技術としては、「高品質話速変換型受聴システム」NHK放送技術研究所（日本音響学会聴覚研究会資料H-92-54：1992年11月20日）に記載されているものがある。

【0004】

【発明が解決しようとする課題】ところで、人間が会話をする場合には、自己が発する声を聞きながら次の声の発声タイミングを計るなど、自己の発する声を聴取することによるフィードバック処理を無意識に行っているものである。

【0005】ところが、上記従来の話速変換機能を有する補聴器では、装用者自身が発した声も話速変換処理をされたのち、装用者に聴取される構成になっているため、話速変換機能を有する補聴器を装着している者は、話速変換処理により遅延された自己の声を聴いて次の発声にフィードバックをかけることになる。しかし、遅延された自分が発した声を聴いて次の発声にフィードバックをかけると、正常に発声を行うことができなくなるといった問題点があった。

10 【0006】この発明の目的は、取り込んだ音声が発用者が発した声かどうかを識別して、装用者が発した声であるときには話速変換処理を行わないで出力する補聴器を提供することにある。

【0007】

【課題を解決するための手段】請求項1の発明は、取り込んだ音声信号を時間軸上で伸長する話速変換手段を備えた補聴器において、装用者の音声信号の特徴情報を記憶する特徴記憶手段と、該特徴情報に基づいて取り込んだ音声信号が発用者の発した音声であるかどうかを識別する音声識別手段と、前記音声識別手段で装用者の音声であることを識別したときには、前記話速変換手段を禁止する話速変換禁止手段と、を備えたことを特徴とする。

20 【0008】請求項2の発明は、取り込んだ音声信号を時間軸上で伸長する話速変換手段を備えた補聴器において、装用者の音声信号を入力し、該音声信号の特徴情報を抽出する特徴抽出手段と、抽出された特徴情報を記憶する特徴記憶手段と、該特徴情報に基づいて取り込んだ音声信号が発用者の発した音声であるかどうかを識別する音声識別手段と、前記音声識別手段で装用者の音声であることを識別したときには、前記話速変換手段を禁止する話速変換禁止手段と、を備えたことを特徴とする。

30 【0009】この発明の話速変換機能を有する補聴器においては、装用者の音声信号の特徴情報を特徴記憶手段に記憶している。請求項2の発明では、装用者の音声信号を入力してその音声信号の特徴情報を抽出し、これを特徴記憶手段に記憶する。

40 【0010】そして、取り込んだ音声信号をこの特徴情報に基づいて判断することにより、装用者が発した声であるかどうかを識別する。この識別において、取り込んだ音声が発用者の発した声であると識別すると、この取り込んだ音声信号に対して話速変換処理を行わずに出力する。

50 【0011】したがって、補聴器は取り込んだ音声信号が発用者が発した声である時には話速変換処理を行わずにこの音声信号を出力し、取り込んだ音声信号が発用者が発した声でない時には話速変換処理を行ってこの音声信号を出力する。よって、補聴器の装用者は、話速変換されていない自分の発した声を聴いて次の発声にフィードバックをかけるので、次の発声を正常に行うことがで

きる。また、話速変換処理された他人が発した声を聴くことになるので、これをゆっくり、はっきり聴くことができる。

【0012】

【発明の実施の形態】図面を参照してこの発明の実施形態である話速変換機能を有する補聴器の構成について説明する。図1～図5は、話速変換機能を有する補聴器において、自己発話時に話速変換を禁止する技術を説明する図である。

【0013】図1は、話速変換機能を有する補聴器の構成を示すブロック図である。この補聴器は、マイクロホン1から入力された音声信号が装用者が発した声の信号であるか否かを判断し、装用者が発した声の場合には、増幅のみして出力し、装用者以外の者が発した声の場合にはこの音声信号を時間的に伸長して出力することにより装用者に聞きやすくしている。話声の信号を時間的に伸長する場合、単に均一に伸長するのみでは自然な話声にすることができない。これは人がゆっくり話す場合、発音の間隔すなわち無音区間が長くなるとともに母音の発音時間が多少長くなるが、子音の発音時間は殆ど変化

しないからである。したがって、この補聴器では、無音区間と母音区間（有声音区間）は伸長し、子音区間（無声音区間）は伸長せずに出力することで自然な話声の伸長を実現している。なお、無音区間と母音区間はそれぞれ個別に伸長する倍率を設定することができる。

【0014】図1において、補聴器は、マイクロホン1、アンチエリアシングフィルタ2、A/D変換回路3からなる入力部、DSP回路12、メモリ回路13からなる信号処理部6、D/A変換回路7、ローパスフィルタ8、イコライザ9、アンプ10、レシーバ11からなる出力部、クロックジェネレータ4、サンプリングクロック発生回路5からなるクロック回路、および、マイクロホン1で受信した音声信号が装用者が発した声であるかどうかを識別する話者識別部14からなっている。

【0015】マイクロホン1は空気振動を音声信号として取り込みこれをアナログの電気信号に変換して出力する。アンチエリアシングフィルタ2は、いわゆるローパスフィルタであり、A/D変換回路3のA/D変換にともなう歪みの発生を防止するため、前記マイクロホン1から入力されるアナログ信号からA/D変換回路3のサンプリング周波数（ $f = 16 \text{ kHz}$ ）の1/2である8

kHz以上の高音域をカットする。A/D変換回路3は、アンチエリアシングフィルタ2で高音域をカットされたアナログ信号をデジタル信号に変換する。A/D変換回路3のサンプリング周波数は上述したように16

kHzであり、この周波数のサンプリングクロックはサンプリングクロック発生回路5から与えられる。サンプリングクロック発生回路5は、クロックジェネレータ4が発生する高速のクロック信号（数十MHz）を分周して上記サンプリングクロックを生成する。クロックジェネレータ4が発生する高速のクロック信号はDSP回路12の動作クロックとして供給される。

【0016】信号処理部6は、前記A/D変換回路3から入力されたデジタル信号に対して伸長処理等の話速変換処理を行う回路であり、マイクロプログラムにより音声信号の切り出しや切り出された音声信号のうち有声音部を伸長するなどの処理を行う。D/A変換回路7は、信号処理部6で処理されたデジタル信号をアナログ信号に変換する。ローパスフィルタ8は、8kHzのカットオフ周波数を有しD/A変換回路7の出力したアナログ信号から折り返しノイズを除去する。このアナログ信号はイコライザ9に入力され、中高音域の信号レベルをブーストするなど装用者の聴力に合わせて周波数特性を変換される。イコライザ9から出力された信号はアンプ10で十分なパワーの信号に増幅されたのちレシーバ11に入力される。なお、イコライザ9を設けずにこの処理をDSP回路12で行うようにしてもよい。レシーバ11は、装用者の耳に装着されるものであり、入力されたアナログ信号を空気振動に変換して装用者の外耳道に放出する。

【0017】マイクロホン1から音声信号が入力されると、この信号はアンチエリアシングフィルタ2を介してA/D変換回路3でデジタル信号に変換されDSP回路12に入力される。同時にこの音声信号が装用者自身の発した声であるか否かが話者識別部14で識別される。話者識別部14が装用者の発した声でない、すなわち、装用者以外の者の発した声であると識別した場合には、DSP回路12は入力されたデジタル信号の時間的伸長を行う。この伸長処理については図2のフローチャートの説明において詳述する。伸長されたデジタル信号はD/A変換回路7でアナログ信号に変換され、ローパスフィルタ8、イコライザ9、アンプ10で折り返しノイズの除去、イコライジング、増幅が施されたのちレシーバ11から放音される。

【0018】一方、話者識別部14が入力された音声信号が装用者自身の発した声であると判断した場合には、DSP回路12は入力されたデジタル信号の伸長を行わずそのままD/A変換回路7に出力する。これにより、装用者自身の発した声はイコライジングおよび増幅はされるが、時間的な伸長は行わずにレシーバ11から放音されることになる。

【0019】図2は、同補聴器の信号処理を示すフローチャートである。マイクロホン1で取り込まれた音声信号は、アンチエリアシングフィルタ2で高音域をカットされA/D変換回路3でデジタル信号に変換される（n1）。このデジタル信号は、DSP回路12を介して順次メモリ回路13に取り込まれる（n2）。メモリ回路13に1フレームとして定められたN個のデータを記憶するとともに、話者識別部14で話者が装用者であるかどうかを識別する（n3）。ここで話者が装用者

でないと識別すると、メモリ回路13に記憶しているデータに対して話速変換処理を行い(n4~n12)、話速変換処理を行ったデジタルデータをD/A変換回路7に出力する。一方、話者が装用者であると識別すると話速変換処理を行わず、入力されたデジタル信号をD/A変換回路7に出力する。

【0020】ここで、n4~n12の話速変換処理について説明する。まず、このフレームが音声区間であるか無音区間であるかを判断する。このため、メモリ回路13に記憶されているフレームデータ(N個のサンプリングデータ)の短時間平均音声パワーE1を求める(n4)。短時間平均音声パワーE1とは、フレーム区間の各データを2乗した値の総和をN分の1した値である。この短時間平均音声パワーE1と予め設定されている閾値P1とを比較し、 $E1 > P1$ のとき音声区間と判定し $E1 < P1$ のとき無音区間と判定する(n5)。

【0021】n5で音声区間であると判定すると、このデータの周波数fzを算出する(n6)。周波数fzはこのフレームデータの零交差回数から抽出する。そして、この周波数の分析を行って有声音または無声音どちらであるかを識別する(n7)。ここで、有声音は主として母音であり、無声音は主として子音である。このため、有声音に含まれる周波数と無声音に含まれる周波数は、おおむね無声音が高く有声音が低い。したがって、零交差回数が少ない場合(周波数が低い場合)には有声音であり、零交差回数が多い場合(周波数が高い場合)は無声音であると識別することができる。この実施形態では、零交差回数から求めた周波数fzが予め設定されている閾値周波数fz1と比較し、 $fz > fz1$ のとき無音区間と識別し、 $fz < fz1$ のとき有音区間と識別するようにした。

【0022】n7で有音区間であると識別された場合、フレームデータを適当な窓幅で切りだし、自己相関関数を求める(n8)。そして、この求めた自己相関関数からピッチ周期を求め(n9)、区間の開始点に最も近い正の微分係数を持つ零交差点を開始点とし、ピッチ区間長後の次のピッチ区間の開始点の一つ前のデータを終了点として伸長用波形を切り出す(n10)。そして、次のデータとして切りだし開始点からピッチ区間長分のデータを接続することで、波形を繰り返す(n11)。一方、n7で無音区間であると識別した場合にはn8~n11の処理を行わずに出力する。すなわち、母音区間に対しては予め設定されている倍率で伸長処理を行い、子音区間に対しては伸長処理を行わない。

【0023】また、n5で無音区間であると識別した場合には予め設定されている長さの無音区間を挿入する(n12)。

【0024】このように処理することにより、補聴器は、他人が発した声を話速変換処理して出力し、装用者が発した声を話速変換処理しないで出力する。よって、

補聴器の装用者は、次の発声を正常に行うことができるとともに、他人が発した声に対してはゆっくり、はっきり聴くことができる。

【0025】ここで、n3での話者が装用者であるかどうかを識別する処理を詳細に説明する。

【0026】最初に、固体伝搬音を利用して話者が装用者であるかどうかを識別する方法を説明する。

【0027】図3は、固体伝搬音を利用して話者を識別する話者識別部の構成を示す図である。話者識別部14は、加速度センサや圧力センサ等の直接接触している物体の振動や機械的变化を検出する骨伝導マイク21と、骨伝導マイク21の出力を増幅する増幅器22と、比較データを記憶する比較データ記憶回路23と、前記増幅器22の出力と前記比較データとを比較する比較器24とを備えている。

【0028】骨伝導マイク21は、補聴器の耳穴に挿入される部分に装着されている。装用者に装着した状態の時に、骨伝導マイク21は装用者の外耳道内壁に直接または保持する構造物を通じて接触するように取り付けられている。骨伝導マイク21は装用者の発声によって生じる外耳道内壁の振動を検出して、これを電気信号に変換して出力する。この信号は、増幅器22によって増幅される。この増幅器22の増幅率は、予め設定されている。この増幅器22で増幅された信号と、比較データ記憶回路23に記憶している比較データとが比較器24で比較される。比較器24は、この比較で増幅器22で増幅された信号が比較データよりも大きい時には、話者が装用者であることを示す信号を出力し、その他の場合には話者が装用者でないことを示す信号を出力する。信号処理部6は、入力された信号が話者が装用者である旨の信号であった場合、マイクロホン1で取り込んだ音声信号に対して話速変換処理をせずそのままD/A変換回路7に出力する。

【0029】図4は空間伝搬音から話者が装用者であるかどうかを識別する構成を示す図である。この回路では、装用者の発する声を検出するための話者識別マイク31を用いている。話者識別マイク31は、例えば、イヤホン型で装用者の耳に挿入して装用者の声を大きく検出するようにされているが、他の音声であっても大きい音声であればこのマイクによって検出される。これを装用者の声と誤認しないようにこの回路は以下の構成になっている。

【0030】図4において、話者識別マイク31が取り込んだ音声信号は増幅器33で増幅され、マイクロホン1で取り込んだ音声信号は増幅器32で増幅される。増幅器32が出力する信号Bと増幅器33が出力する信号Aとを比較器34で比較する。増幅器32、33の増幅率は、それぞれマイクロホン1、話者識別マイク31を取り付ける位置に応じてそれぞれ個別に決定されている。比較器34は信号Aが信号Bよりも大きい時には、

話者が装用者であることを示す信号を出力する。DSP回路12はこの信号が入力されたときマイクロホン1から取り込んだ音声信号に対する話速変換処理を禁止する。

【0031】以上の構成により、装用者以外が発した音声が大きくなり、話者識別マイク31が大きく取り込んでも、これを装用者の声と誤認することがなくなる。

【0032】図4は、アナログ回路で信号Aと信号Bとを比較しているが、DSP回路12の機能を用いてこの比較を行うこともできる。この場合には、話者識別マイク31の出力信号をA/D変換してDSP回路12に入力するとともに、図5に示す動作を行う。まず、話者識別マイク31で受信した信号Aとマイクロホン1で取り込んだ信号Bを所定の係数で増幅したのち(n21、n22)、増幅された信号Aとを比較する(n23)。信号A、Bに対する係数はそれぞれマイクロホン1、話者識別マイク31を取り付ける位置に応じてそれぞれ個別に決定されている。比較の結果信号Aが信号Bよりも大きい時には、話者が装用者であると判断し(n24)、その他の場合には話者が装用者でないと判断する(n25)。このようにすることで、装用者以外の人物の発生が極めて大きい場合でも、正確に話者を識別することができる。信号処理部6では、この信号に基づいて上記した話速変換処理を禁止する。

【0033】次に、装用者の発する声の特徴を記憶して、入力された音声信号の特徴からこの信号が装用者の声であるか否かを判断するこの発明の実施形態である話速変換機能を有する補聴器について説明する。

【0034】図6は同補聴器のブロック図である。この実施形態において図1に示した補聴器と構成において異なる点は、話者識別部14を備えていない点およびメモリ回路13に装用者が発した音声の特徴を記憶する装用者音声特徴記憶エリア13aを備えている点で異なる。

【0035】また、この補聴器の動作において上記実施形態と異なる点は、話者が装用者であるかどうかを識別する処理が異なるだけで、その他の処理(話速変換処理等)については同一である。ここでは、話者が装用者であるかどうかを識別する処理のみ説明し、その他の動作については説明を省略する。

【0036】図7は、この実施形態の話者識別処理を示すフローチャートである。補聴器の装用者は事前に自分が発した声の特徴を装用者音声特徴記憶エリア13aに記憶させている(図7(A)参照)。この、装用者音声特徴記憶エリア13aに自分が発した声の特徴を記憶させるには、補聴器を登録モードに切り換えてマイクロホン1に装用者が自分の声を取り込ませる(n31)。この音声信号はアンチエイリアシングフィルタ2を通過し、A/D変換回路3でデジタル信号に変換されて信号処理部6に送られる。信号処理部6では、DSP回路12がこの信号の線形予測分析を行い(n32)、装用者で

ある人物の発した音声の特徴を抽出する(n33)。そして、この抽出した音声の特徴を装用者音声特徴記憶エリア13aに記憶する(n34)。

【0037】通常モードの時(使用状態の時)には(図7(B)参照)、補聴器はマイクロホン1で音声信号を取り込むと(n41)、この取り込んだ音声信号を上記と同じ経路で信号処理部6に入力する。信号処理部6では、DSP回路12がこの信号の線形予測分析を行い(n42)、この音声信号の特徴を抽出する(n43)。そして、装用者音声特徴記憶エリア13aに記憶している音声の特徴と、n43で抽出した音声の特徴とを比較し、類似点を評価する。この評価における評価点が予め設定されている値よりも大きい場合には、話者が装用者であると識別し、評価点が予め設定されている値よりも小さい場合には、話者が装用者でないと識別する。

【0038】なお、話者識別部14をスイッチで構成し、装用者自身に識別信号を入力させるようにしてもよい。例えば、信号処理部6を該スイッチがオンしている間話速変換処理を禁止する構成にすれば、装用者は自分が話している間だけこのスイッチをオンすることによりこの発明の機能を実現することができる。また逆に、信号処理部6をスイッチがオフしている間話速変換処理を禁止する構成にした場合には、装用者は自分が話している間だけこのスイッチをオフすればよい。なお、このスイッチは、補聴器の動作を停止させるためのメインスイッチではなく、話速変換処理を許可/禁止するスイッチであり、話速変換処理が禁止されている間でもマイクロホン1から入力された音声信号は増幅されてレシーバ11から出力される。

【0039】以上説明したように、補聴器は取り込んだ音声信号が装用者が発した声である時には話速変換処理を行わずにこの音声信号を出力し、取り込んだ音声信号が装用者が発した声でない時には話速変換処理を行ってこの音声信号を出力する。よって、補聴器の装用者は、話速変換されていない自分の発した声を聴いて次の発生にフィードバックをかけるので、次の発声を正常に行うことができる。また、話速変換処理された他人が発した声を聴くことになるので、これをゆっくり、はっきり聴くことができる。

【0040】なお、請求項2の発明を上記実施形態に対応させて具体化すると以下になる。音声識別手段は、装用者の発した声を線形分析して特徴を抽出して記憶する装用者音声記憶手段と、取り込んだ音声信号を線形分析して特徴を抽出する特徴抽出手段と、前記特徴抽出手段で抽出した特徴と前記装用者音声記憶手段に記憶している特徴を比較して類似点から取り込んだ信号が装用者が発した音声であるかどうかを識別する手段である。

【0041】

【発明の効果】以上のように、この発明によれば、装用者の音声信号の特徴に基づいて装用者の発声を識別することができるため、話者識別マイクなどを別途装用しなくても、装用者の発声を識別し、装用者が発した声には話速変換を行わず、他人が発した声に対しては話速変換を行う。したがって、補聴器の装用者は、話速変換されていない自分の発した声を聴くことができるので、次の発声を正常に行うことができる。また、他人が発した声をゆっくり、はっきり聴くことができる。

【図面の簡単な説明】

【図1】話速変換機能を有する補聴器の構成を示すブロック図

【図2】同補聴器の信号処理を示すフローチャート

【図3】固体伝搬音を利用して話者を識別する話者識別*

* 部の構成を示す図

【図4】空気伝搬音を利用して話者を識別する話者識別部の構成を示す図

【図5】空気伝搬音を利用して話者を識別する処理を示すフローチャート

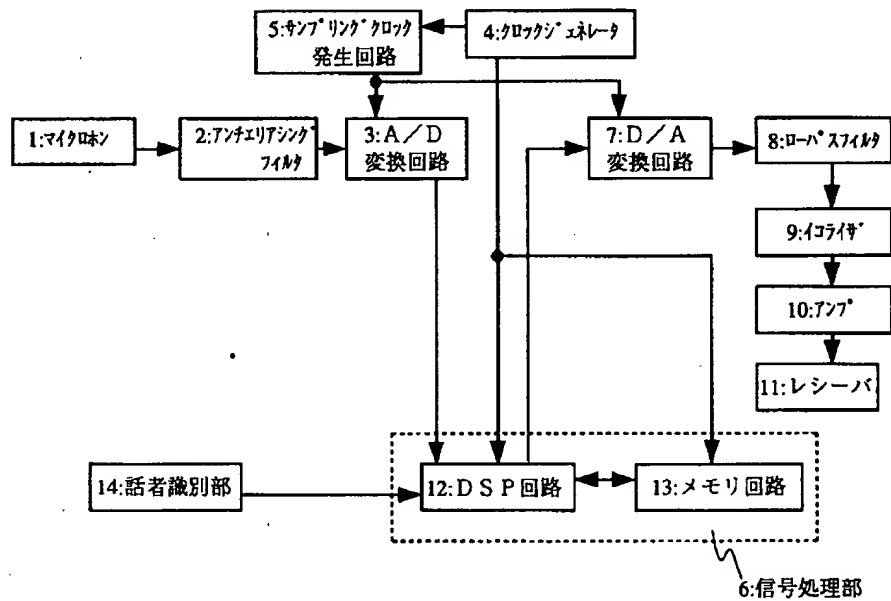
【図6】この発明の実施形態である話速変換機能を有する補聴器の構成を示すブロック図

【図7】この発明の実施形態の話者識別処理を示すフローチャート

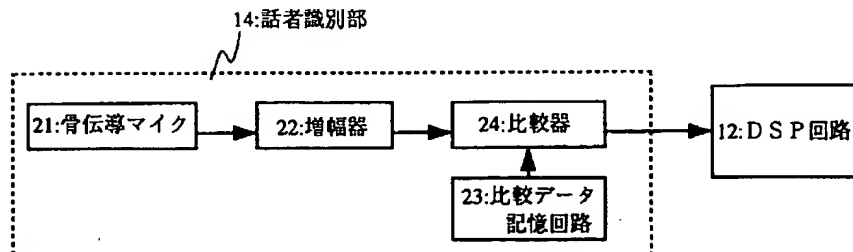
10 【符号の説明】

1…マイクロホン、6…信号処理部、12…DSP回路、13…メモリ回路、13a…装用者音声特徴記憶エリア

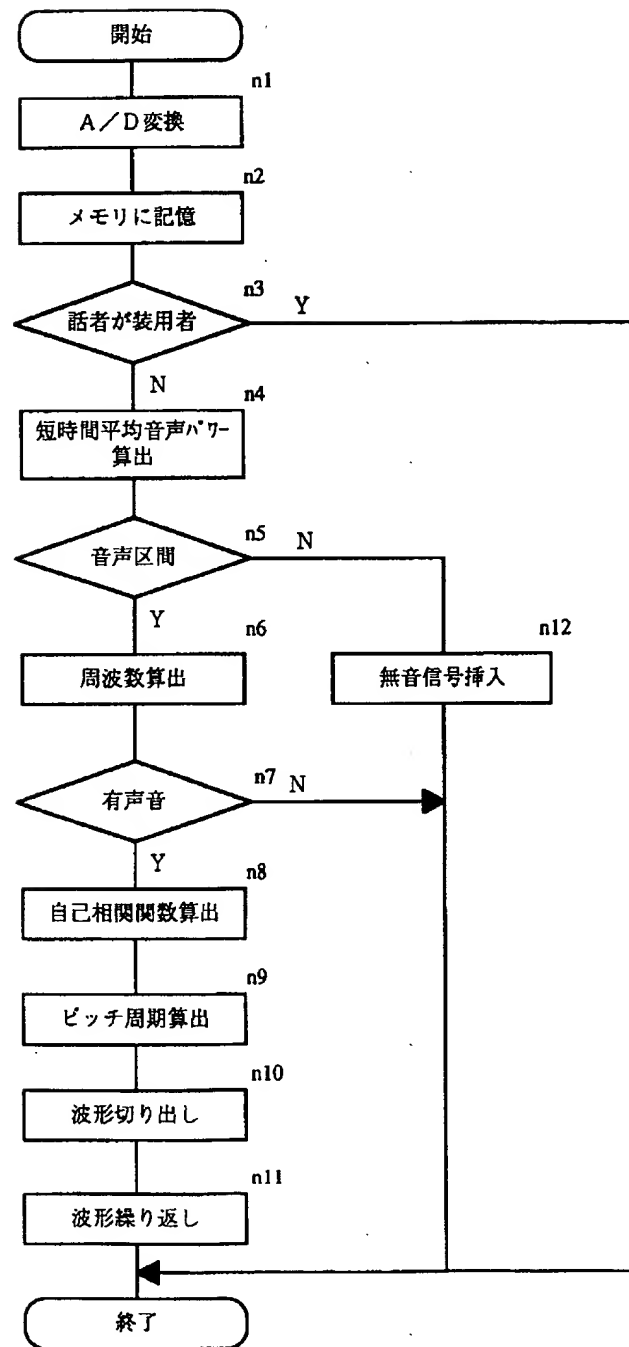
【図1】



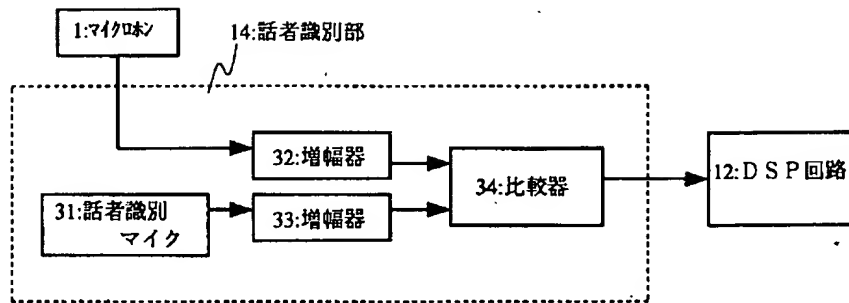
【図3】



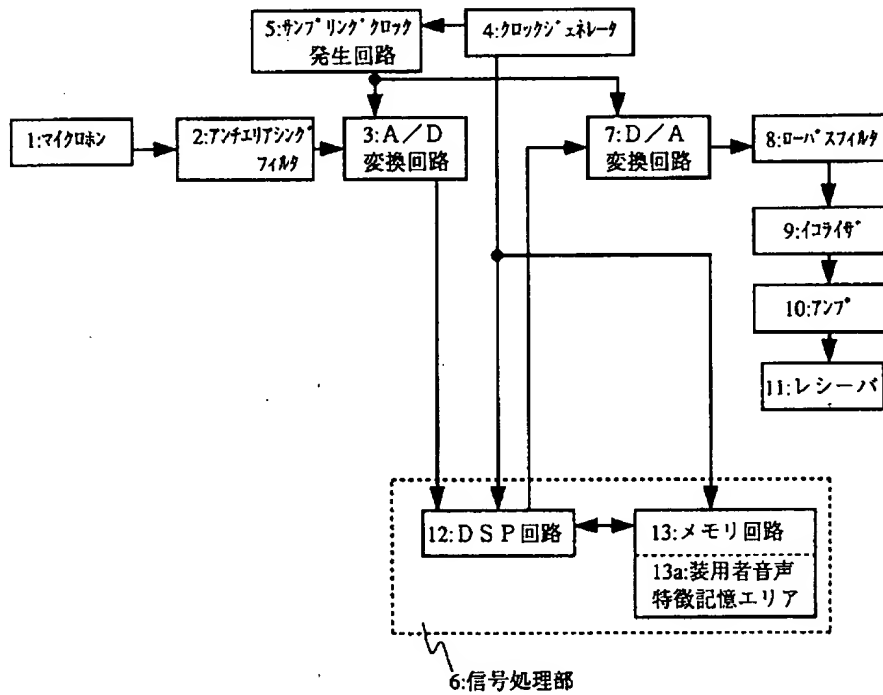
【図2】



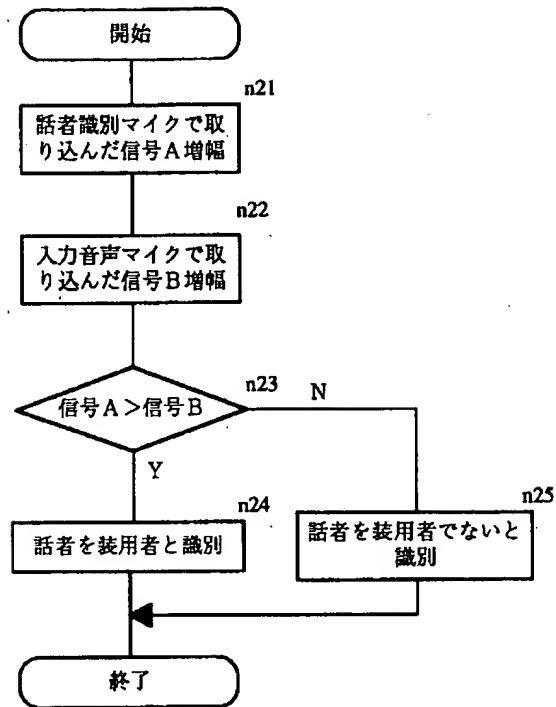
【図4】



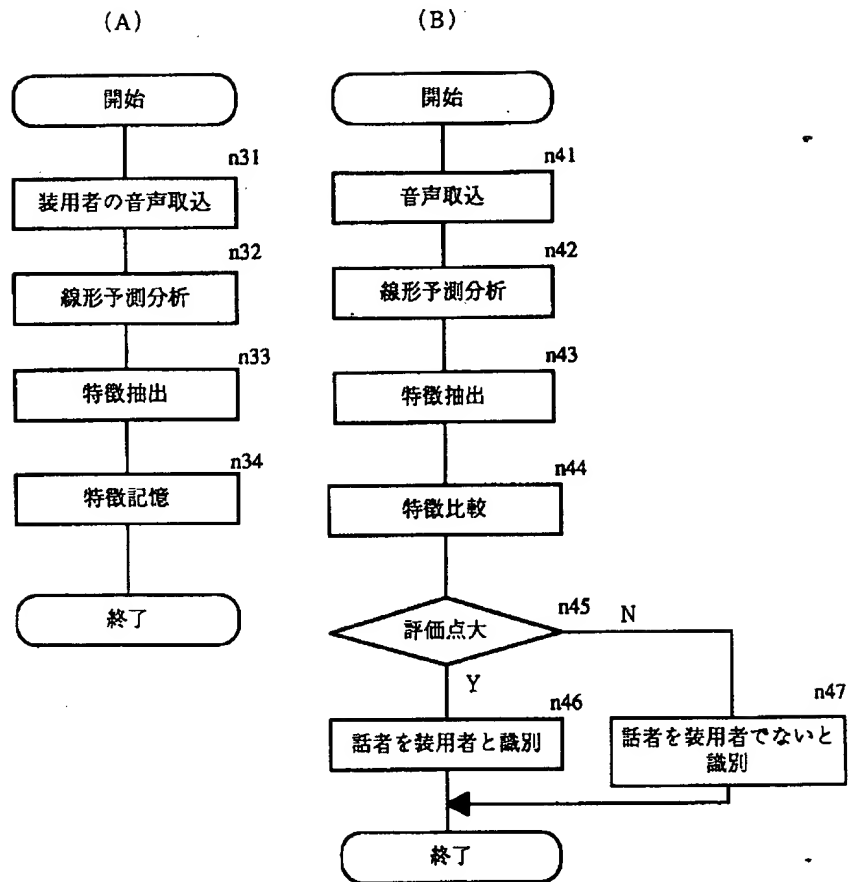
【図6】



【図5】



【図7】



JP-H10-290498

[Publication Number]

Japanese Patent Laid-Open No. 10-290498

[Publication Date] (1998) October 27

[Title of the Invention]

Hearing aid having a speed of speech conversion facility

[Application Number] Japanese Patent Application No. 10-116809

[Indication of Division]

Partitioning of Japanese Patent Application No. 6-249244

Filing date (1994) October 14

[Applicants] [Identification Number] 000004075

[Name] YAMAHA Inventor: Katsufumi Kondo, Kouji Tanidaka

[Embodiment for carrying out the invention]

Configuration of "hearing aid having a speed of speech conversion facility" which is embodiment is described when taken with the drawing.

FIG. 1-FIG. 5 is a chart explaining technique to prohibit speed of speech conversion in self-speech production in hearing aid having a speed of speech conversion facility. FIG.1 is a block diagram which shows configuration of hearing aid having a speed of speech conversion facility. This hearing aid determines it whether audio signal input from mic 1 is in signal of "the voice that user uttered". In the event of the voice that user emitted, it is output with status only for amplification. This audio signal is extended in terms of time, and, in the event of the voice that a person aside from user emitted, it is output. By this, it is easy to be heard in user, and it is done. When signal of speech sound is extended in terms of time, it seems to become the following. Extending it cannot merely make it natural speech sound uniformly.

When a person talks slowly, "interval of articulation(silent interval)" gets longer. And pronunciation time of a vowel sound becomes a little long.

However, most of the pronunciation time of a consonant does not vary.

Thus, with this hearing aid, silent interval and vowel sound interval (phonic interval of vocal sound) extend it, and it is output.

In addition, consonant interval (unvoiced sound interval) outputs it without extending it. By this, this realizes elongation of natural speech sound.

In addition, silent interval and vowel sound interval can set "scale factor to

extend individually". In FIG. 1, hearing aid comprises the thing that the following says it. Mic 1, Anti-aliasing filter 2, Input unit comprising A/D converting circuit 3, DSP circuitry 12, Signal processing division 6 comprising memory circuit 13, D/A conversion circuitry 7, Lowpass filter 8, Equalizer 9, Amp 10, Output part comprising receiver 11, Clock generator 4, Clock circuit comprising sampling clock generation circuit 5, Speaker identification department 14 (it identifies it whether the audio signal which it received with mic 1 is in the voice that user emitted). It comprises these. Mic 1 takes in air vibration as audio signal.

This is converted into analog electrical signal, and it is output.

Anti-aliasing filter 2 is so-called lowpass filter.

A/D converting circuit 3 is warped with AD conversion.

It prevents this generate. Therefore, from an analog signal input from mic 1, the following is cut.

Sampling frequency ($f=16\text{kHz}$) of A/D converting circuit 3, 1/2 "high frequency sound of 8kHz greater or equal."

This is cut.

A/D converting circuit 3 converts "an analog signal cut high frequency sound with anti-aliasing filter 2" to a digital signal.

Sampling frequency of A/D converting circuit 3 is 16kHz as had stated above.

A sampling clock of this frequency is given from sampling clock generation circuit 5.

As for sampling clock generation circuit 5, divide does "high-speed clock signal (several decades mega Hz) that clock generator 4 occurs".

And it generates a sampling clock.

The high-speed clock signal that clock generator occurs is provided as an action clock of DSP circuitry 12.

Signal processing division 6 is a circuitry doing a speed of speech conversion process (a tensile process) for "digital signal input from A/D converting circuit 3".

By a microprogram, the following processes are done.

It cuts audio signal. And it extends voice part among cut audio signal.

D/A conversion circuitry 7 converts "a digital signal processed with signal processing division 6" to an analog signal.

These processes are done in a microprogram.

Lowpass filter 8 comprises cut off frequency of 8kHz.

It turns around "the analog signal which D/A conversion circuitry 7 output", and noise is removed. This analog signal is input into equalizer 9.

Thus a signal level of middle-high frequency signal level area can leave boost.

A frequency characteristic is changed to hearing acuity of user as above.

Signal output from equalizer 9 is amplified in amp 10 by signal of enough power.

And it is input into receiver 11.

This process may be done with DSP circuitry 12 without providing equalizer 9.

Receiver 11 is loaded by ear of user.

An input analog signal is converted into air vibration, and it is ejected in external acoustic meatus of user.

When audio signal is input from mic 1, this signal goes through anti-aliasing filter 2, and it is converted to a digital signal with A/D converting circuit 3.

And it is input into DSP circuitry 11.

With speaker identification region 14, it is identified at the same time whether this audio signal is the voice that user oneself emitted.

When it was identified if there is not speaker identification department 14 in the voice that user emitted (it is the voice that a person aside from user emitted), DSP circuitry 12 extends a time-variant of an input digital signal. It is explained in detail in description of flowchart of FIG. 2 about this tensile process.

An extended digital signal is converted into an analog signal with D/A conversion circuitry 7.

After, in "9, lowpass filter8, equalizer9, amp 10", "destruction of lapel noise, equalizing, amplification" were put, sound is output from receiver 11.

When it was determined that input audio signal is the voice that user oneself gave off by speaker identification department 14, it seems to become the following.

DSP circuitry 12 does not extend an input digital signal.

It just outputs it to D/A conversion circuitry 7. As for the voice that user oneself emitted, equalizing and the amplification are done by this.

As for the time elongation, it is output sound from receiver 11 without doing it.

FIG.2 is flowchart which shows signal conditioning of the hearing aid.

As for the audio signal taken in with mic 1, it is cut high frequency sound with

anti-aliasing filter 2.

And it is converted into a digital signal with A/D converting circuit 3 (n1).

This digital signal is taken in memory circuit 13 through DSP circuitry 12 sequentially (n2).

To memory circuit 13, "data of an N piece brought peace to as one frame" is memorized.

It is identified with speaker identification department 14 therewith whether a speaker is in user (n3).

When it is identified here if there is not a speaker in user, it seems to become the following.

It does a speed of speech conversion process as against data memorizing to memory circuit 13. (n4-n12).

The digital data which did a speed of speech conversion process is output to D/A conversion circuitry 7.

On the other hand, a speed of speech conversion process is not done when it is identified when a speaker is user.

It outputs an input digital signal to D/A conversion circuitry 7.

Speed of speech conversion handling of n4-n12 is described now.

At first this frame determines voice interval or silent interval.

Because of this short time mean voice power E1 of memorized frame data (sampling data of an N piece) is demanded from memory circuit 13 (n4).

"Short time mean voice power E1" is the value that did summation of "the value that square did each data in frame interval" in $1/N$.

"Threshold value P1 set beforehand" is compared with "this short time mean voice power E1".

For the case $E1 > P1$, it is determined to being voice interval. For the case $E1 < P1$, it is determined to being silent interval (n5).

It is determined that it is voice interval, and it is (n5)

By this, frequency fz of this data is calculated (n6).

Frequency fz extracts it from a zero intersection number of times of this frame data:

And this frequency is analyzed, and identification (voice or unvoiced sound) is done (n7).

Voice is mainly a vowel sound, and, here, unvoiced sound is mainly a consonant.

Because of this, in "frequency included in voice" and "frequency included in unvoiced sound", unvoiced sound is almost high, and voice is low.

Thus, it is voice when there are a little zero intersection number of timeses (when frequency is low).

When there are much zero intersection number of timeses (when frequency is high), it is unvoiced sound.

In this embodiment, frequency fz pursued from a zero intersection number of times compares it with threshold value frequency $fz1$ set beforehand.

In there, it is identified as follows.

($fz > fz1$, unvoiced sound interval):

($fz < fz1$, voice interval):

When when it was voice interval in $n7$, was identified, frame data begin to be cut in suitable window spreading. And an auto correlation function is demanded ($n8$).

And pitch period is pursued from this found autocorrelation function. ($n9$) a zero crossing having "the regular differential coefficient which is the nearest to origin of interval" is done with origin.

There is origin of the next pitch interval of "back of overall length of pitch interval".

It assumes data before this one end point. And it cuts and brings down a waveform for elongation ($n10$).

And it begins to drain it off as the next data. It connects data for overall length of pitch interval from origin. By it, it repeats a waveform ($n11$).

When it identified it when it is unvoiced sound interval in $n7$, it does not handle $n8$ - $n11$. And it outputs it.

In other words, for vowel sound interval, elongation is processed with scale factor set beforehand.

For consonant interval, a tensile process is not done.

When it was identified when it is silent interval in $n5$, silent interval of overall length set beforehand is inserted ($n12$).

As thus described, as for the hearing aid, speed of speech conversion processes the voice that another person emitted by handling it, and it is output.

And the voice that user emitted is output without handling speed of speech conversion.

Therefore, user of hearing aid can do the next utterance normally.

In addition, it gets possible to listen for the voice that another person emitted clearly slowly.

A process to identify here whether a speaker in $n3$ is in user is explained in detail.

Solid-borne sound is used first.

And a method to identify whether a speaker is in user is explained.

FIG. 3 is a chart to show configuration of speaker identification department to identify a speaker in using structure borne sound. Speaker identification department 14 comprises a thing in the following.

Bone conduction microphone 21 detecting "vibration, variation of machine" of the object which "acceleration sensor, pressure sensor" come in contact with directly.

Amplifier 22 amplifying output of bone conduction microphone 21.

Comparison data memory circuit 23 storing comparison data.

Comparator 24 which compares the comparison data with output of amplifier 22. 14 comprises these.

Bone conduction microphone 21 is loaded by "part interposed in ear canal" of hearing aid.

At the time of status attached to user to form a specified shape, bone conduction microphone 21 is fitted in external acoustic meatus intine of user.

It is fitted to come into contact directly.

It is fitted through holding structure to come into contact.

Bone conduction microphone 21 detects vibration of external acoustic meatus intine occurring by means of phonation of user.

This is converted into electrical signal, and it is output.

This signal is amplified by means of amplifier 22.

Amplification factor of this amplifier 22 is set beforehand.

"Comparison data stored to comparison data memory circuit 23" is compared with "signal amplified with this amplifier 22" with comparator 24.

In this comparison, 24 outputs "signal showing that a speaker is user". In that case, when signal amplified with amplifier 22 is bigger than comparison data.

In the event of others, "the signal which a speaker shows a few thing in in user" is output.

Signal processing division 6 is handled as follows.

When input signal was the signal which showed that a speaker was user, for audio signal taken with mic 1, speed of speech conversion is not processed.

And it is just output to D/A conversion circuitry 7.

FIG. 4 is a chart to show configuration to identify whether a speaker is in user in from space propagation sound.

In this circuit, "speaker identification microphone 31 to detect the voice that user emits" are used. For example, it is inserted into ear of user with

earphone type, and it is big, and speaker identification microphone 31 detects a voice of user. It is other voices, and a big voice is detected with this microphone. This circuitry is equal to or less than it not to mistake this for a voice of user, is arranged.

In FIG. 4, the audio signal which speaker identification microphone 31 took in is amplified with amplifier 33.

The audio signal which was busy with mic 1 is amplified with amplifier 32.

Signal A which amplifier 33 outputs is compared with signal B which amplifier 32 outputs with comparator 34.

Amplifier 32, 33 amplification factor accept position installing "mic 1, speaker identification microphone 31", and it is determined individually.

When signal A is bigger than signal B, comparator 34 shows that a speaker is user, it supplies a signal.

When this signal was input, DSP circuitry 12 prohibits a speed of speech conversion process for the audio signal which was busy from mic 1.

By the above-mentioned configuration, an uttered voice aside from user grows big.

Or speaker identification microphone 31 takes it in greatly.

Even if there are these, it disappears to mistake this for a voice of user.

FIG. 4 compares signal B with signal A in analog circuit.

This comparison can be done by means of a facility of DSP circuitry 12.

For this case, A/D converts an output signal of speaker identification microphone 31 to being similar.

And it is input to DSP circuitry 12.

Along with these processes, action shown in FIG. 5 is done.

Signal B taken in with signal A and mic 1 received with speaker identification microphone 31 is amplified with a prescribed coefficient (n21, n22).

And amplified signal A is compared with them (n23).

A coefficient as opposed to signal A, B accepts position installing "mic 1, speaker identification microphone 31", and it is determined individually.

When signal A is bigger than signal B, it is determined that a speaker is user as a result of comparison (n24).

In the event of others, a speaker judges it not to be it in user (n25).

As thus described a speaker can be identified precisely by doing it when generate of a person aside from user is extremely big.

It is signal processing division 6, and, based on this signal, the described

above speed of speech conversion process is prohibited.

And the vocal special feature that user emits is memorized.

It is determined whether this signaling is in a voice of user from a special feature of input audio signal.

Embodiment of device to have these special features toward is explained (hearing aid having a speed of speech conversion facility)

FIG. 6 is a block diagram of the hearing aid.

In this embodiment, a period to be different from hearing aid shown in FIG. 1 in configuration is equal to or less than it. Speaker identification department 14 is not comprised.

"User voice special feature storage area 13a" storing the phonic special feature that user emitted to memory circuit 13 is had.

It is different from it by possessing this.

A process to identify whether a speaker is in user is different from the embodiment in a different thing in action of this hearing aid.

It is same about other processes (speed of speech conversion processes).

Here, only a process to identify whether a speaker is in user is explained.

Description is omitted about other action.

FIG.7 is flowchart which shows speaker identification handling of this embodiment.

User of hearing aid makes "user voice special feature storage area 13a" store the vocal special feature that oneself gave off beforehand (cf. FIG. 7 (A)).

Hearing aid is changed in register mode to make this 13a store the vocal special feature that oneself gave off.

And user inputs one's voice into mic 1 (n31).

This audio signal passes anti-aliasing filter 2.

With A/D converting circuit 3, it is converted into a digital signal.

And it is sent to signal processing division 6.

With signal processing division 6, DSP circuitry 12 does linear predictive coding of this signal (n32).

"The phonic special feature that the person who is user gave off" is extracted (n33).

And this extracted phonic character is stored in user voice special feature storage area 13a (n34).

Time (time of busy condition) (FIG. 7 (B) cf. FIG. 7 (B)) of a normal mode.

As this occurs, hearing aid takes in audio signal with mic 1 (n41).

This busy audio signal is input into signal processing division 6 in a path the

same as the above.

With signal processing division 6, DSP circuitry 12 does linear predictive coding of this signal (n42).

A special feature of this audio signal is extracted (n43).

The phonic special feature that extracted it in n43 is compared with a phonic special feature memorized in 13a.

And a similarity is evaluated.

When a validation period in this validation is bigger than set point, it is identified when a speaker is user.

When a validation period is smaller than set point, it is identified if there is not a speaker in user.

Speaker identification department 14 is arranged with a switch, it may make recognition signal input in user oneself.

By way of example only, while a switch turns on signal processing division 6, it is done in configuration to prohibit a speed of speech conversion process.

While oneself talks, user turns on this switch.

By this, a facility of this devise can be realized.

In addition, while a switch turns off signal processing division 6 adversely, it is done in configuration to prohibit a speed of speech conversion process.

User should turn off this switch then while oneself talks.

In addition, this electric switch is not main switch to stop action of hearing aid.

This is the switch that authorization / prohibits a speed of speech conversion process.

While a speed of speech conversion process is prohibited, audio signal input from mic 1 is amplified, and it is output from receiver 11.

As discussed above it seems to become the following.

Busy audio signal is the voice that user emitted.

As for being similar, hearing aid does not do a speed of speech conversion process then.

And this audio signal is output.

Busy audio signal is not the voice that user emitted.

At this time, a speed of speech conversion process is done.

And this audio signal is output.

Therefore, user of hearing aid listens to the voice that oneself who cannot leave speed of speech conversion gave off.

And user of hearing aid submits feedback to the next utterance.

Therefore, the next utterance can be done normally.

In addition, the voice that speed of speech conversion processed another person emitted is heard.

By this, this can be listened to clearly slowly.

When it makes devise of claim 2 support the embodiment, and it realizes, it seems to become the following.

A voice discriminating means has user phonetic memory means.

Linear analyzes the voice that user emitted into the means, and a special feature is extracted, and it is stored.

"A special feature extraction means" linear analyzes the audio signal that it was busy, and to extract a special feature is comprised.

A special feature stored in the user phonetic memory means is compared with character extracted in the feature extraction means.

It is means to identify whether the signal which was busy from the similarity is in the voice that user emitted.

Brief Description of the Drawings

[FIG. 1]

The block diagram which shows configuration of hearing aid having a speed of speech conversion facility

[FIG. 2]

Flowchart to show signal conditioning of the hearing aid

[FIG. 3]

Block diagram of speaker identification department to identify a speaker using sound transmitted through a solid

[FIG. 4]

Block diagram of speaker identification department to identify a speaker using sound transmitted through air

[FIG. 5]

Flowchart to show a process to identify a speaker using sound transmitted through air

[FIG. 6]

A block diagram to show configuration of hearing aid having the speed of speech conversion facility that is embodiment of this devise in

[FIG. 7]

Flowchart to show a process of the speaker identification which is embodiment of this devise

[Denotation of Reference Numerals]

- (1) mic
- (6) signal processing division
- (12) DSP circuitry
- (13) memory circuit
- (13a) area to store a user voice special feature